


Article

# An Empirical Study of a Trustworthy Cloud Common Data Model Using Decentralized Identifiers

Yunhee Kang <sup>1</sup>, Jaehyuk Cho <sup>2,\*</sup> and Young B. Park <sup>3</sup><sup>1</sup> Division of Computer Engineering, Baekseok University, Cheonan 31065, Korea; yhkang@bu.ac.kr<sup>2</sup> Department of Electronic Engineering, Soongsil University, Seoul 06978, Korea<sup>3</sup> Department of Software Science, Dankook University, Yongin 16891, Korea; ybpark@dankook.ac.kr

\* Correspondence: chojh@ssu.ac.kr

**Featured Application:** The proposed DID-based service model is designed as an agent that is based on a platform with DID. It provides interoperability, privacy, and efficiency to manage identity in cloud CDM.

**Abstract:** The Conventional Cloud Common Data Model (CDM) uses a centralized method of user identification and credentials. This needs to be solved in a decentralized way because there are limitations in interoperability such as closed identity management and identity leakage. In this paper, we propose a DID (Decentralized Identifier)-based cloud CDM that allows researchers to securely store medical research information by authenticating their identity and to access the CDM reliably. The proposed service model is used to provide the credential of the researcher in the process of creating and accessing CDM data in the designed secure cloud. This model is designed on a DID-based user-centric identification system to support the research of enrolled researchers in a cloud CDM environment involving multiple hospitals and laboratories. The prototype of the designed model is an extension of the encrypted CDM delivery method using DID and provides an identification system by limiting the use cases of CDM data by researchers registered in cloud CDM. Prototypes built for agent-based proof of concept (PoC) are leveraged to enhance security for researcher use of ophthalmic CDM data. For this, the CDM ID schema and ID definition are described by issuing IDs of CDM providers and CDM agents, limiting the IDs of researchers who are CDM users. The proposed method is to provide a framework for integrated and efficient data access control policy management. It provides strong security and ensures both the integrity and availability of CDM data.

**Keywords:** common data model; collaborative work; identity; distributed ledger; credential; decentralized identifiers



**Citation:** Kang, Y.; Cho, J.; Park, Y.B. An Empirical Study of a Trustworthy Cloud Common Data Model Using Decentralized Identifiers. *Appl. Sci.* **2021**, *11*, 8984. <https://doi.org/10.3390/app11198984>

Academic Editors: Seongsu Cho and Bhanu Shrestha

Received: 17 August 2021

Accepted: 24 September 2021

Published: 27 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Today, the key issue of medical services is moving forward from treatment to prevention and management of diseases [1]. Medical institutions and companies have been promoting technology development in related fields to provide services based on artificial intelligence and big data technology using medical data [2–4]. Clinical studies based on patient data from numerous hospitals can provide more meaningful results. However, since each hospital uses a different structure of Hospital Information System (HIS), the need for a CDM is recognized for systematic data management and integrated research [5]. CDM is a data structure defined to efficiently utilize hospitals' data. It is composed based on international standard terms and has different components depending on the purpose. Through CDM, various data structures and meanings for each institution are converted to have the same structure and meaning, and various difficulties caused by different data structures between institutions can be solved when conducting multi-institutional joint research.

However, despite the advantage of being able to efficiently manage data, it still has problems such as regulation and protection of personal information due to the fundamental characteristics of medical data. In the existing CDM, identity management methods have mainly been isolated, centralized, and federated. These methods have limitations in interoperability due to closed identity management, Identifier (ID) leakage, and subordination with external ID management subjects [6–8]. In cloud CDM, it is necessary to design a secure cloud on a permission-type block chain in which the access control of the authorized and registered researcher is established [9].

In order to use the CDM data, the request of access permission from the researcher and Institutional Review Board (IRB) approval are required in the data supervision process, and the results of the process are maintained in the block chain. When various hospitals and research institutes take part in cloud CDM, an access control system is required to prove the researcher's permission to participate in the research as well as the interoperability of the participating institution's systems. In the operating organization of cloud CDM, a stepwise qualification process is required according to the roles of CDM provider, CDM consumer, researcher, and IRB. In the cloud CDM environment, verifiable identities are essential to handle CDM data securely and ensure the system supports the reliable and tamper-evident nature of the subject's identity. It allows the development of independent digital identities rooted on a distributed ledger [10,11]. It also helps bring building applications with a solid digital foundation of trust by enabling the verifiable credentials model. For identity, verifiable credentials are derived from a registry.

Due to restrictions of the domestic medical law, sharing medical information outside the medical institutions in a domestic medical information utilization environment is restricted except when the patient himself/herself requests his/her own records for personal information. Because the data management system is fragmented and centralized, the exchange and use of medical information is limited, and the information management is insufficient, making it difficult for cooperative research [12,13].

One of important points about data sharing in this regulatory aspect is the IRB. For clinical studies of medical data, researchers must comply with the conditions set forth in the Research Participation Regulations. In the cloud CDM environment, the researcher has special requirements that the researcher's affiliated institution may be different from the CDM data provider. To solve this problem, the researcher must obtain permission to participate in the research from the IRB of the institution that provides clinical information and controls the conduct of the research.

This paper describes the application of decentralized identifiers (DID) to prove user identity in the cloud CDM environment. DIDs' transactions are configured using Hyperledger Indy, and CDM subjects are configured as agents based on Hyperledger Aries [14,15] to evaluate the behavior of CDM use cases. Here, we design and prototype a DID-based user-centric identification system to support the research of registered researchers in the cloud CDM environment involving multiple hospitals and research institutions. The prototype is an extension of the delivery method of encrypted CDM using DID and provides the identification system by limiting the use case of the CDM data of the researcher registered in the cloud CDM. The prototype constructed for agent-based PoC (proof of concept) is utilized for enhanced security of researcher use of ophthalmic CDM data. In this paper, the CDM identity schema and its definition are described by limiting the identity of main entities.

This proposed method aspires to provide a unified and efficient data access control policy management framework. It provides strong security and ensures both the integrity and the availability of CDM data. It aims to build upon and improve existing data governance processes between different organizations, translating the information sharing policies they already apply in their current operational interactions into electronically enforceable rules embedded in credentials.

The main contributions of our work can be summarized as follows:

1. DID-based user-centric identification is the first to approach supporting researchers autonomously with the identity verification with a verified proof without a third party having central authority in the cloud CDM.
2. We propose and solve the service model that extends the DID basic model in order to solve the structural problem where it is difficult to participate in external researchers in the hospital situation related to IRB approval.
3. We validate user access control by applying the DID service model in the safe data transfer process between hospitals in Korea.
4. Our service model provides high interoperability by operating the prototype of identity proof using the standard messaging environment using DIDComm.

## 2. Related Works

### 2.1. CDM

With the widespread adoption of electronic health records (EHRs) in healthcare systems, clinical data are entering the digital era [16]. Large-scale EHR data analysis has produced influential discoveries, which have enabled the practice of precision healthcare [17]. However, there are many barriers that limit the usefulness of EHR data, primarily revolving around available expertise. Since EHR data are large and typically stored in relational databases, many clinical experts and scientists have no experience, lack sufficient time to spare, and need knowledge of Structured Query Language (SQL) programming. Moreover, the structure and data components of an EHR system are complex and require strong familiarity to be utilized most effectively. Many people solve this problem by building effective collaboration across multiple disciplines (e.g., doctors working with data science teams) but enabling more researchers to directly work with data is important. Thus, CDM facilitates the interoperability of EHR data for research, and it requires strong familiarity to enable many researchers to handle the data directly [18].

CDM is a typical database model of medical information standardization for clinical data-based research [19]. Simultaneous multi-center analysis can be performed in the form of standardized schemas and vocabulary systems and has been continuously updated considering existing limitations. This allows us to transform different data structures and meanings from an institution to have the same structure and meaning, thus solving the difficulties due to difference in data structures for each institute [20].

There are various CDMs such as Observational Medical Outcomes Partnership (OMOP)-CDM, Sentinel-CDM, and national-scale clinical research network (PCORnet) CDM [21]. In particular, the OMOP CDM is a common data model developed and operated by the Observational Health Data Sciences and Informatics (OHDSI) international consortium, with more than 200 organizations from 14 countries participating in the transition to CDM [22]. OMOP-CDM uses a common medical terminology system called the OMOP code as well as the same data structure, enabling an integrated analysis of clinical healthcare databases across multiple institutions. A CDM database built in each institution has the advantage of being able to perform more efficient and systematic analysis using the already developed CDM-based open-source standard analysis methods and analysis programs from libraries and web bases [23].

### 2.2. Blockchain and Its Application in CDM

Blockchain is a technology in which a number of transaction details are bundled to form a block. Additionally, several blocks are connected like a chain using hashes, and then a number of people copy and store them in a distributed manner [24,25]. It allows anyone to make reliable and secure transactions. Blockchain can not only be used for cryptocurrency but also for all data processing that has online transaction history and requires history management. Blockchain-based smart contracts, logistics management systems, document management systems, medical information management systems, copyright management systems, social media management systems, game item management systems, electronic voting systems, identity verification systems, etc., can be used in various ways [25].

Although the demand for the use of medical data is increasing, medical information contains personal information, so there are restrictions on its use. In the domestic health and medical field, as the need to provide medical services tailored to patient characteristics by integrating genomic information, treatment, clinical information, and lifestyle information is increasing, it is essential to secure a security system for the safe use of medical information.

Blockchain has a function that can be used for safe and clean data distribution of CDM data in a collaborative research environment in which multiple institutions participate. In a collaborative research environment, CDM providers and consumers operate blockchain nodes and manage the process of transactions related to access to CDM data by researchers.

SimplyVital [26] uses a private blockchain network to share patients' personal medical data with multiple medical research institutions but maintains it on the distributed ledger of the patient's medical information provider. However, maintaining medical information on the blockchain is an illegal matter of domestic medical information, and there is a limitation to applying it to the joint use of medical information. OmniPHR [27] operates based on a blockchain to convert and maintain the dispersed personal medical information of patients into a standardized data format and manage the authority to access the data from any device. As access control is performed, there is a limitation that access control must be performed for each individual researcher.

In this paper, we present a study on blockchain technology for user identity management of medical data and a model for cross-institutional CDM data access control. The proposed model is based on a decentralized identity to provide self-sovereignty, and through this, proves the qualifications of researchers in an environment in which multiple institutions participate. The DID model for credentialing is currently an effective approach for accessing CDM data, a unique use case in the medical field.

### 2.3. Decentralized Identifier (DID)

From a security standpoint, an identity is an entity such as a user, virtual group, or organization that can be used to define permissions on a security item. The two main functions of an identity are accountability and access control. The identifier is used to uniquely identify entities and give unique names to data to express its characteristics.

DID is a technology that allows individuals to have complete control over their information, unlike the existing identification method controlled by a central system. By using DID, if an individual interacts with a specific institution, the owner of the identity information can control whether or not the information is provided so that the identity information can be managed transparently. We classified the ID management techniques required for handling the CDM between cooperative organizations into four types and compared their characteristics [28,29]. These management types are isolated, centralized, federated, and self-sovereign.

In the isolated type, the identity of the user is managed by service, and the user has to go through the self-authentication (membership registration) and identity authentication (authentication) procedures for each service. This is to securely establish and operate identities within a single institution. However, it is not suitable for the secure service operation of multi-participant cloud CDM and requires significant costs for user authentication and access control [30].

The centralized identity management is a method that centrally manages the user's identity in terms of efficiency, and the construction and operation of the identity management system are superior to the isolated type based on individual identity management. When users register their IDs in the central management system, they can access and use various services through the central identity management server. The centralized type with these technical characteristics is suitable for centrally managed and controlled single cloud CDM operation. However, if a failure of a single central management system happens, a single point of failure that cannot use the entire service is unavailable. It also has limitations in terms of scalability and interoperability.

In the federated type, different service providers form a trust relationship and jointly manage the user's identity [29]. Hospitals participating in the cloud CDM network can operate by applying standards such as SAML. However, federated identity management is required to establish a trust relationship between the hospital and the cloud CDM first. This has the constraint that it has to depend on a specific service provider through ID management, and there is also a single point of failure problem.

In the self-sovereign identity (SSI) type, individual information can be controlled by the individual himself/herself and is based on Distributed Ledger Technology (DLT) without the intervention of a third party [30]. The identity information required for the service can be selectively submitted through the channel, and the reliability of the submitted information can also be proven without the intervention of a third party [22,23]. A representative example is DID, which is being standardized by the W3C. Individuals are independent of any single organization because they provide their identity as their identity provider. A self-sovereign identity system can use blockchain to look up distributed identifiers without a central directory. When you register your ID in the blockchain, your ID proof based on the block chain is issued and you sign the ID proof. When a general establishment presents their blockchain ID proof, the establishment verifies that the user's identity information is appropriate by inquiring about their ID to the blockchain.

DIDs are a hashed form of a public key. The private keys for DIDs are stored in a wallet. The wallet is used for allowing any user to store their digital information securely on a personal device [29–31]. An agent is any application that stores and uses DIDs. It is the software that interacts with other entities via DIDs. The verifiable DID model consists of three roles, issuer, holder, and/or verifier. Figure 1 presents a basic DID model proposed by the World Wide Web Consortium (W3C). The statements of verifiable credentials are generated by an issuer. The presentations based on the credential are sent to the verifier to attest the authenticity of verifiable credentials issued by the issuer.

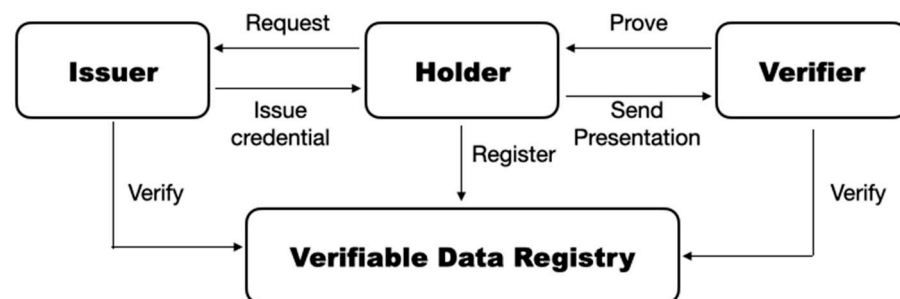


Figure 1. The DID model proposed by W3C.

Hyperledger Indy is a public and permissioned blockchain platform tailored to build DID. The following describes the main characteristics of Indy:

- It provides individuals with independent control over their personal data.
- It has to allow interoperability with other decentralized ledgers.
- It supports the attribute and claims schema system written to the ledger for dynamic discovery of claim types.

Hyperledger Aries provides a library for handling verifiable digital credentials. The envelope of the messages between agents has been standardized in the form of the DIDComm protocol. DIDComm describes how messages should be encrypted and decrypted in transport. The agent is an entity working in the cloud CDM, where it interprets messages on behalf of its organization and executes a command to support secure access to the CDM service. The agent has secure storage that is used for all the information collected by it. In this paper, we design a researcher as one of agents in the cloud CDM.

In summary, in terms of information security, identity management plays an important role in preventing illegal access from the outside. However, the traditional identity management model relies on a third-party central system for information management.



In this approach, it is difficult for the central system to have complete reliability, and it involves the problem of information exposure to the outside. Identities must not be held by a single third-party, even if it is a trusted entity temporarily [32,33]. Additionally, the central system has a single fault risk [34,35]. To solve the above problems, the proposed model is based on SSI, where individual researchers maintain their identity, and supports secure CDM data transmission and solves the privacy problem.

The proposed model is used for safe CDM data transmission and access control of transmitted data in the cloud CDM environment. Table 1 compares the traditional identity management technology and the DID in the following five aspects.

**Table 1.** Comparison of traditional identity managements and DID.

	Traditional Identity Management			DID
	Isolated	Centralized	Federated	
Privacy and protection	Low	Low	Low	High
User control and consent	Low	Low	Moderate	High
Dependency	Moderate	High	High	Low
Fault tolerance	High	High	Moderate	Low
Usability	Low	Low	Moderate	High

- Privacy and protection

The rights of users must be protected on the set of tasks such as handling data. The users must be able to choose their privacy model. When personal data are disclosed, that disclosure implicates the minimum amount of information required to complete the given task.

- User control and consent

In the cloud CDM, researchers must control their identities. They may enable referring to their own identity, updating it, and accessing their own data. Their identities must not be held by a single third-party entity in the cloud CDM.

- Dependency

Each of the organizations is running an independent corporation without dependency in the cloud CDM.

- Fault tolerance

To access a system handling user identity, it enables the continued working of the system despite failures or malfunctions in cloud CDM.

- Usability

The access right is a system granted to users according to the domain policy. The researcher's experience must be consistent with their expectations in a research process.

### 3. The Extended to Identify Management Scheme for Cloud CDM

#### 3.1. The Cloud CDM Model

In order to collect and integrate clinical data of multiple hospitals, it is required to solve the heterogeneity of data structure and format, differences in quality and quantity of data, technical limitations of interoperability, and security issues. CDM should support the linking of common analysis codes for electronic medical record (EMR) resource linkage to support integrated data analysis of research institutions, without leaking sensitive personal information.

Data extracted from EMRs tend to be stored in different relational database schemas. Figure 2 illustrates the conventional concept of CDM and its operation scheme derived from several sources of EMR in hospitals.

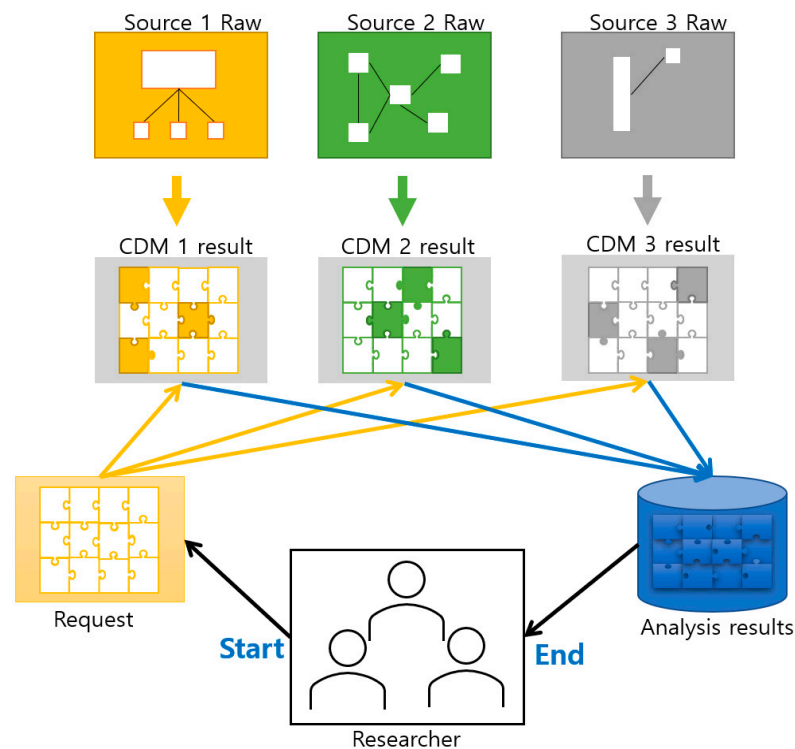


Figure 2. Conventional concept of the Common Data Model (CDM) and operation scheme.

The cloud CDM reference model shown in Figure 3 is a partial result from the previous works and consists of several CDM providers and CDM consumers participating [10]. Using this presented reference model, clinical researchers can isolate and securely distribute CDM data.

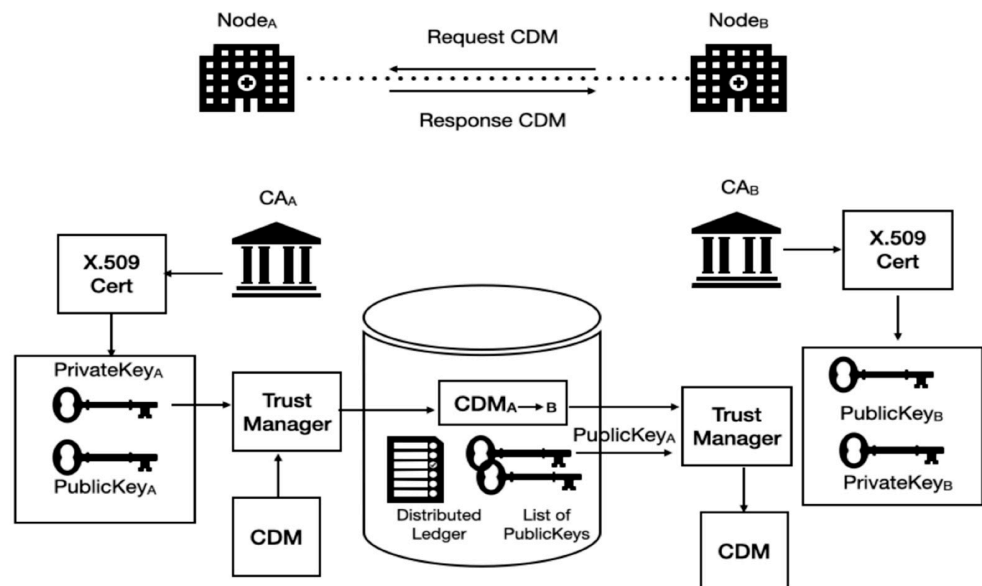


Figure 3. Concept of the Secure-Cloud Common Data Model (SC-CDM).

- Cryptography can be used for protecting information, using a hash value to maintain management of large-capacity CDMs. Encryption can be used to protect information using symmetric and asymmetric keys to maintain the management of large-capacity CDMs.
- A distributed ledger is used to provide data integrity and share information through a CDM signature.

- In the process of data creation and use, the distributed ledger guarantees data integrity, and transparently signed CDM can be accessed.

### 3.2. The Operation Scheme for Trustworthiness in CDM Cloud

In this paper we are focused on how to guarantee the trustworthiness using DID among the entities in cloud CDM. Hence, this model has no consideration of authentication and authorization based on in-person and group verification of cloud CDM. In cloud CDM, it is necessary to design a secure cloud on a permission-type blockchain in which the access control of authorized and registered researcher is established. In order to use the CDM data, the request of access permission from the researcher and IRB approval are required in the data supervision process, and the results of the process are maintained in the blockchain.

The following shows the process for uploading the CDM derived from the researcher's query in Figure 4:

1. A researcher registered in a medical institution, Hospital B, sends a query to the EMR DB managed Hospital A.
2. The researcher requests the trust manager of Hospital A for CDM to hold the cloud CDM based on the result of the query.
3. The trust manager of Hospital A obtains the IRB's approval for the request for the EMR data with the credential for identifying the researcher.
4. The trust manager in Hospital A builds the approved EMR data into CDM data and its metadata associated with encryption keys and storing the CDM encrypted to distribute to a repository in cloud CDM.
5. The trust manager in Hospital A uploads the encrypted data to the cloud CDM.

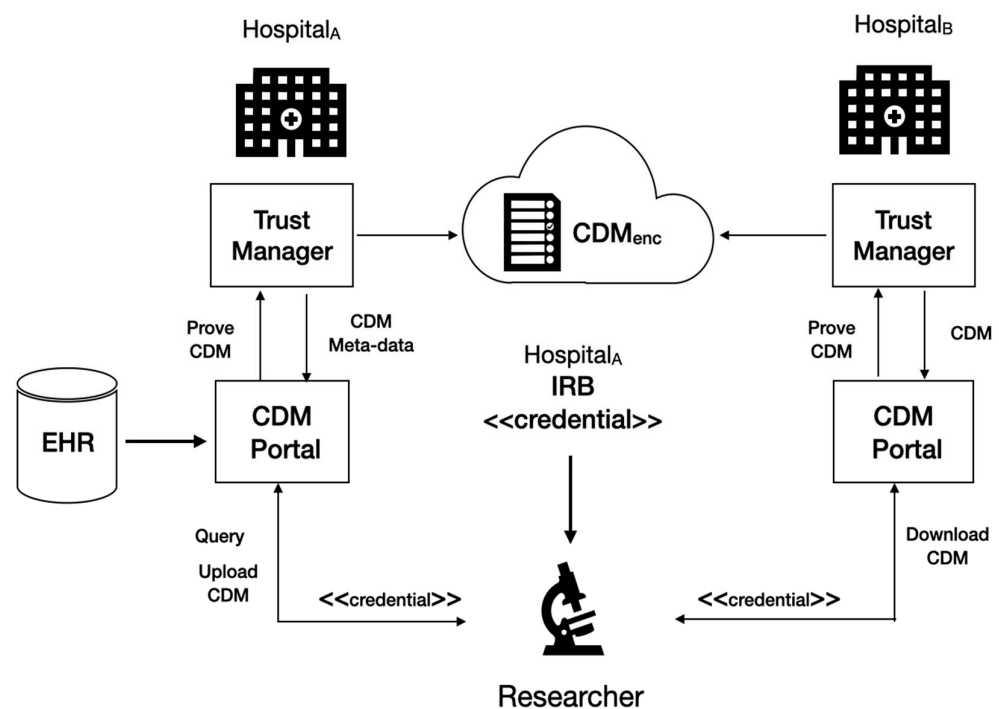


Figure 4. The overall concept of authentication and authorization in cloud CDM.

We assume the requirements of authentication and authorization as the research background. The authentication is the basic process of verifying that the entities (researcher, IRB, CDM provider, CDM consumer) are who they claim to be before allowing access. In the context of cloud CDM, authorization determines the entitlement of an entity to perform tasks that are authorized within the system. A user's authorization and authentication are initially activated by an identity provider (IRB) and provide CDM data about the person granted by the IRB.



### 3.3. The Basic DID Model for Cloud CDM

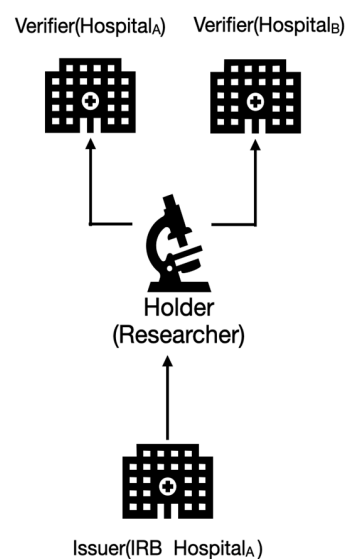
In the basic model, the identity information necessary for the information subject to receiving the desired service from the verification agency is issued and submitted by the personalization agency. To ensure the validity of the identity information issued by the personalization agent, the certificate of the personalization agent is stored in a verifiable data registry. The verification body that has received the proof of identity verifies the proof in the registry and provides services. A credential is an attestation of qualification, competence, or authority issued to an entity (e.g., an individual or organization) by a third party with a relevant or de facto authority or assumed competence to do so.

If research involves human subjects or is regulated by the Food and Drug Administration (FDA), it requires review and approval from an institutional review board (IRB) or the Human Subjects Office. It is the responsibility of all faculty and students to obtain IRB approval or Exempt determination before initiating any human subjects research projects.

Hence, IRB uses a public DID published globally. The IRB play a role as a verifiable credential issuer. Since the researcher as holder of the credential may present the credential to anyone, the identity (via the public DID) of the issuer must be part of what the verifier learns from the presentation. The verifier can investigate (as necessary) to decide if they trust the issuer. The public DID of IRB is put on a blockchain so that it can be globally resolved. It is used to establish secure, point-to-point messaging channels between the agents of the participants. With a verifiable credential, DIDs are used as the identifier for IRB as the issuer in cloud CDM.

IRB (the issuer) DID is used to uniquely identify the issuer and is resolved to obtain a public key related to the DID. That public key is then used to verify that the data in the verifiable credential did indeed come from the issuer. This public DID ensures that the verifier knows who issued the credential a holder presents.

Figure 5 shows the basic DID model for cloud CDM. Node A represents a CDM provider, and Node B represents a CDM consumer. Two trust managers located in the service broker play role as agents of the CDM provider operated in Node A and the CDM consumer operated in Node B for trustily delivering the CDM (represented as  $CDM_{A \rightarrow B}$  in Figure 3). The verifiers may not fully trust the researcher without a verifiable credential (VC) and want to share only a subset of data or respond with data retrieved from a particular query. They might also want to share different subsets of data to the researcher. The grant of access may also need to be revoked, updated, or set to expire.



**Figure 5.** The DID-based trust model for cloud CDM.

In cloud CDM, credentials need to be issued and verified through the following application use cases:

- A researcher is a member of a group of researchers of a specific subject on which he or she wants to conduct research and is assigned a role as a research participant through IRB approval and is registered. Through the IRB, researchers are provided with a certificate of research participation (issuing research participation certificate through IRB).
- CDM users apply to the creation of CDM data, encryption of the generated CDM data, and proof of access service for use in distributed storage.
- CDM users apply for access service verification for decryption and distributed storage of CDM data in the process of accessing the created CDM data.

The following is assumed to operating environment:

- For CDM use, researchers are registered with the CDM provider or user organization. Through the registration process, the researcher assumes that the mutual trust relationship of the cloud CDM participating organizations can be established, managed, and managed through the certificate authority (CA).
- IRB approval documents are used for the purpose of price proof for CDM provision and use (users who have received credentials in the IRB use DID to identify their identity).
- The researcher is provided with the ID of the CDM provider through the approval of the IRB.
- The CDM provider decides to provide the CDM through verification of the researcher’s identity certificate. After qualification verification, the CDM provider performs encryption and distributed storage of CDM data.
- CDM users access the encrypted and distributed CDM data through verification of the researcher’s identity certificate via the CDM consumer.
- The researcher’s research participation certificate maintains the research period as an attribute and allows access to CDM services and data limited to the valid period.

The overall process of issuing and verifying credential when handling CDM in use-cases is shown in Figure 6.

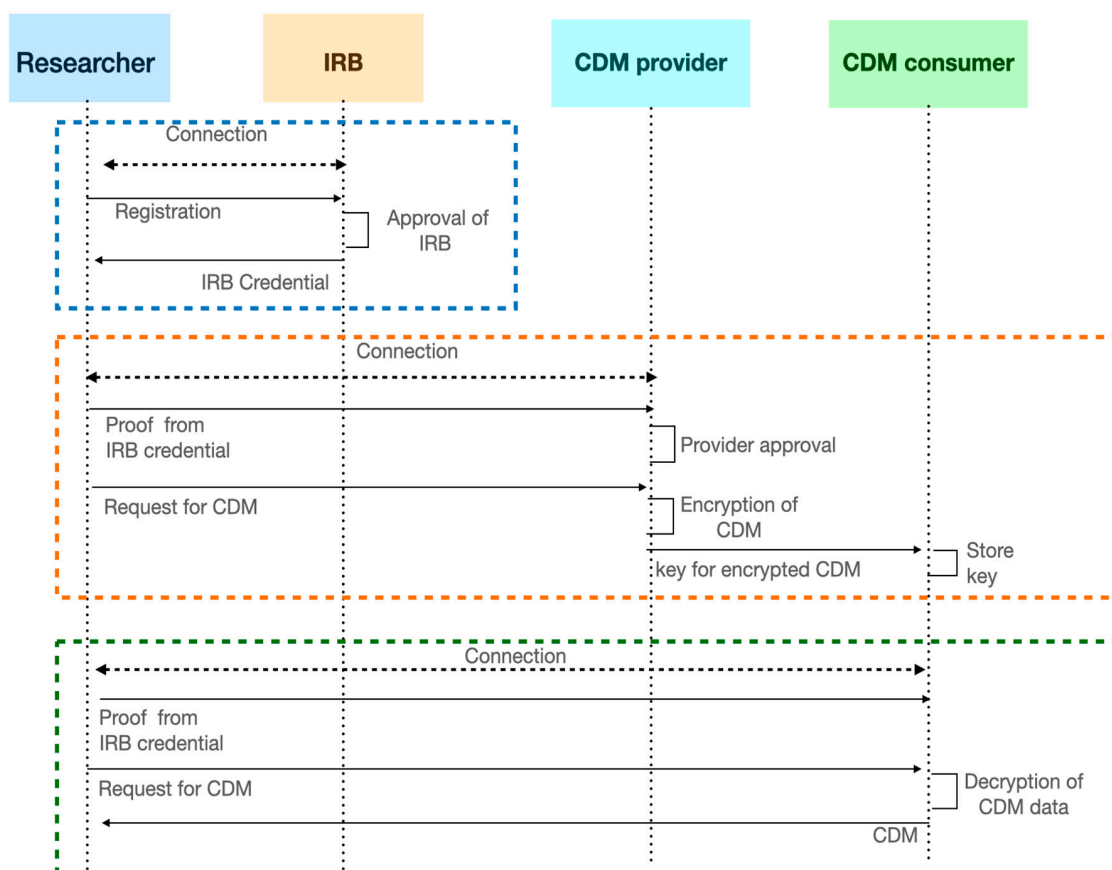


Figure 6. The process of issuing and verifying credential when handling CDM.

### 3.4. Credential Definition of Identity

Self-sovereign identity consists of an identifier and identifier data. In cloud CDM, identifiers use DID, and identifier data consists of several attribute information. The main attribute information for identity consists of personal information, credentials, and verifiable presentation. A legal entity's identity (i.e., an individual or an organization) can be represented using a set of attributes associated with the entity (such as name and role). The identity of the CDM providing and consuming institutions and the participant of these institutions is expressed in various attribute information. Identity management provides the functions for maintaining the identity data and their access control. IRB identity is defined based on its schema. The identity certificate is issued by the IRB provider. Figure 7 is the schema definition for CDM identity stored in Indy DLT.

```
{
  "schema_id": "T8j4DNmf7Us8tTzpvoK6No:2:IRB schema:51.1.53",
  "schema": {
    "ver": "1.0",
    "id": "T8j4DNmf7Us8tTzpvoK6No:2:IRB schema:51.1.53",
    "name": "IRB schema",
    "version": "51.1.53",
    "attrNames": [
      "affiliation",
      "name",
      "GCP",
      "IRB_no",
      "approved_date",
      "role",
      "timestamp"
    ],
    "seqNo": 38
  }
}
```

Figure 7. Schema definition for CDM identity issued by IRB.

The following shows the schema defined for the issued CDM identity stored in the DLT. It shows that the schema was created by the IRB through the credential definition ID.

- Schema ID: T8j4DNmf7Us8tTzpvoK6No:2:IRB schema:51.1.53
- Cred def ID: T8j4DNmf7Us8tTzpvoK6No:3:CL:38:irb.agent.IRB\_schema
- Type: CRED\_DEF
- Reference: 38
- Signature type: CL
- Tag: irb.agent.IRB\_schema
- Attributes: affiliation, approved\_date, gcp, irb\_no, master\_secret, name, role, timestamp

After the IRB agent starts up, the researcher agent establishes a trust channel with the IRB agent, and then the IRB performs DID exchange with the researcher. Algorithm 1 describes the steps for establishing a connection between these agents.

---

#### Algorithm 1 Establishing Trusted Connections

---

- 1: *Researcher* agent exchanges DIDs with the *IRB* agent to establish a DIDComm channel.
  - 2: *IRB* offers an audited researcher credential over this channel.
  - 3: *Researcher* accepts and stores the credential in their wallet.
- † Audited researcher credential is specified by IRB.
- 

### 3.5. Issuing IRB Credential

With a connection with the researcher's agent established the IRB issuer can interact with that agent. It might ask for a presentation to confirm the identity of the researcher.

Eventually, it will reach the point of needing to issue a credential to the researcher. To do that, the controller passes to the framework the type of the credential, the data for the claims, and the connection identifier for the researcher, and the framework (for the most part) takes care of issuing the credential for the given research subject. Note that after offering the credential to the researcher, the response might not come back for hours. This is not an issue, the issuer framework will just wait. Once the credential is issued, an identifier for the credential is given back to the controller, which again stores that with the rest of the information it keeps on the researcher. To issue an Indy credential, the simplest instance of the protocol must have three steps:

- The issuer sends the holder an offer message.
- The holder responds with a request message.
- The issuer completes the exchange by sending the holder an issue message containing the verifiable credential.

The access policy defines programmatically the requirements for authorization to access CDM. The access policy defines these rules based on the CDM, user/group assignments, and ownership assignments. The IRB credential represents the access policy of CDM. Algorithm 2 describes the steps for issuing credential, and the detailed issuing flow is as follows.

1. The holder sends a proposal to the issuer (issuer receives proposal). When the holder starts with sending a proposal, it uses the `/issue-credential-2.0/send-proposal` endpoint.
2. The issuer sends an offer to the holder based on the proposal (holder receives offer). The issuer receives the proposal and can respond with an offer using the `/issue-credential-2.0/records/{id}/send-offer` endpoint. After this offer, the flow continues with the holder responding with a request.
3. The holder sends a request to the issuer (issuer receives request). If the holder automatically accepts offers and turns them into requests, then the issuing of credentials would be completely automated. That improves privacy—making the user in control of when and whom to share information with.
4. The issuer sends credentials to the holder (holder receives credentials). The issue credential protocol is used to enable an issuer to provide a holder with a verifiable credential. In this protocol:
  - There are two participants (issuer, holder).
  - There are four message types (propose, offer, request, and issue).
  - There are four states (proposed, offered, requested, and issued).
5. The holder stores credentials and sends acknowledgement to the issuer. Verifiable credentials are issued to the user and stored in his/her digital wallet, and the user decides when and where to use them.
6. The issuer receives acknowledgement.

---

#### Algorithm 2 Issuing credential

---

- 1: for each *Researcher* agent do
  - 2: Initiate DID Exchange with *CDM provider* agent to establish DIDComm channel.
  - 3: *Researcher* agent delivers the CDM selected to *CDM provider* agent via DIDComm channel.
  - 4: *CDM provider* offers Verified CDM token credential over DIDComm.
  - 5: *Researcher* agent accepts and stores the credential
  - 6: *CDM provider* encrypts the CDM and delivers the cipher CDM to *CDM consumer* agent with the IRB number approved by IRB
  - 7: end for
- † The CDM is derived from the EMR of in CDM provider  
 † Verified CDM token credential is specified by PROVIDER
-

### 3.6. Proof the Credential

Privacy is important when dealing with CDM. The entities using DIDs will be able to express only the portions of their credentials. This expression of a subset of one's credential is called credential presentation. Specifically, the presentation refers to the verifiable data received by a verifier. Instead of typing in the name, address, and government ID, a presentation of that information is provided from verifiable credentials issued from IRB by an authority trusted by the verifiers, CDM provider, and CDM consumer. The verifiers can automatically accept the claims in the presentation (if they trust the issuer) without any further checking.

Instead of obtaining the data directly from the issuer IRB, the data from the issuer comes from the holder, researcher, and the cryptographic material to verify that the authenticity of the data comes from the distributed ledger. This reduces the number of integrations that have to be implemented between issuers and verifiers. A researcher can be issued a professional accreditation credential from the relevant authority (e.g., the College of Physicians and Surgeons) and the claims verified (and trusted) by medical facilities in real time.

Should the doctor lose his or her accreditation, the credential can be revoked, which would be immediately in effect. This would hold true for any credentialed profession, be it lawyers, engineers, nurses, tradespeople, real estate agents, and so on.

## 4. Implementation

### 4.1. Experimental Setup

In this section, the design of the experiments is introduced. Detailed information of our hardware and software configurations is described in Table 2. To run von-network and agents, a docker engine is controlled by those containers. Each of the containers is running as a light-weighted virtual machine.

**Table 2.** Hardware and software configuration.

Item	Model
CPU	Intel(R) Xeon(R) E-2134
RAM	16 Gbyte
OS	Linux 3.1.0
Docker	19.03.8
Docker-compose	1.21.0

Hyperledger Indy node management is permissioned. It has its own ledger and stores/reads public information in the distributed ledger that is reliably elected. The nodes communicate to agree (reach consensus) on what transactions should be written and in what order. To start Hyperledger Indy nodes, a von-network is used. It is a portable development of Hyperledger Indy with a ledger browser. The von-network plays a role as a Hyperledger Indy public ledger sandbox instance. In this work, it is running in docker locally.

Figure 8 shows the von-network with four nodes for identity management in cloud CDM. The von-webserver has a web interface that allows you to browse the transactions in the blockchain.

Before issuing a credential, a credential definition as well as its schema needs to be created. Both the schema and the credential definition are recorded on a von-network. Hyperledger Aries Cloud Agent Python (ACA-Py) is a foundation for building a verifiable credential (VC) ecosystem [35]. It operates in the second (DIDComm Peer to Peer Protocol) and third (Data Exchange Protocols) layers of the Trust Over IP framework using DIDComm messaging and Hyperledger Aries protocols in Figure 9.

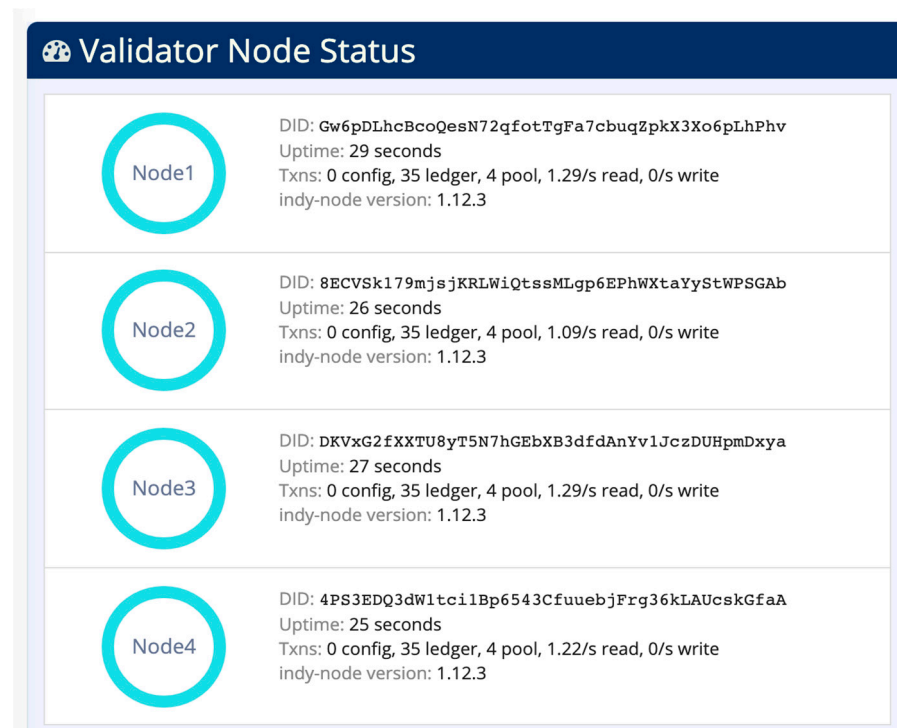
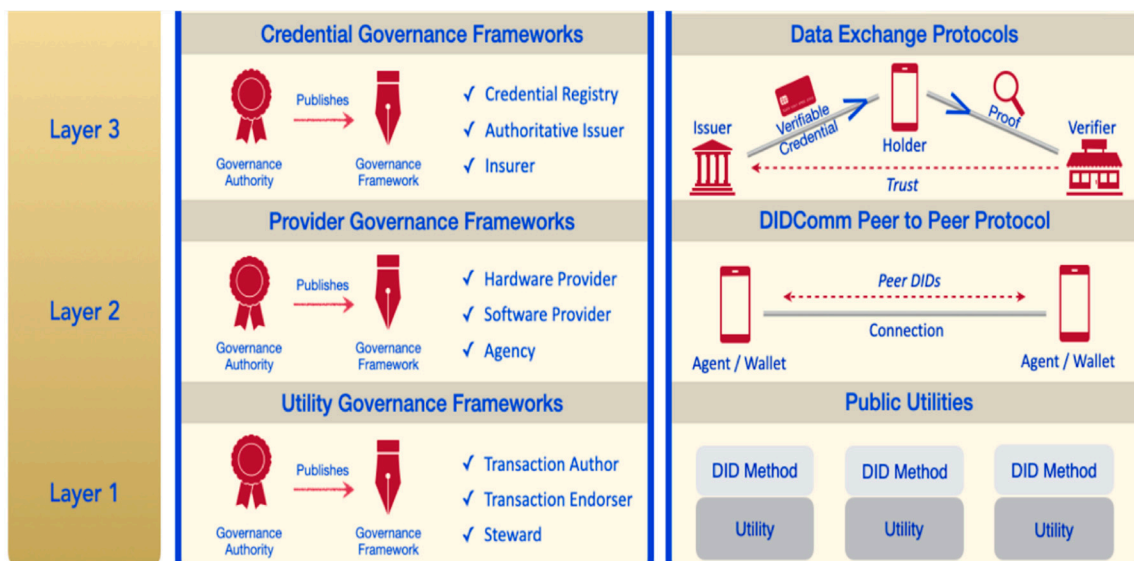


Figure 8. The von-network running in cloud CDM.



(a) Trust over IP governance stack.

(b) Trust over IP technology stack.

Figure 9. Trust over IP framework [36].

A business logic controller is written for the development of a given use case, and the created controller uses the ACA-Py library based on AIP (Aries Interop Profile) 2.0. AIP 2.0 protocols are used for issuing, verifying, and holding VCs that work with a Hyperledger Indy distributed ledger. The von-network is used to represent a credential format named AnonCreds (Anonymous Credentials). It is a kind of detailed implementation of zero-knowledge proof (ZKP) support.

A ZKP is a kind of cryptographic method, and its use in blockchain appears to be promising in cases where existing blockchain technologies can adapt a ZKP to address specific business requirements focusing on data privacy [37]. It proves attributes for an



entity (a person, organization, or thing) without exposing a correlatable identifier about that entity. That claims from verifiable credentials can be selectively disclosed, meaning that just some data elements from credentials, even across credentials can (and should be) provided in a single presentation. By providing them in a single presentation, the verifier knows that all the credentials were issued to the same entity.

Four agents, the researcher, IRB, provider, and consumer are developed. Those agents are written in Python by using ACA-py library. Agents that receive a message from another entity post a webhook internally over HTTP, allowing the controller to respond appropriately. Note that this can include requesting the agent to send further messages in reply. More details can be seen in Table 3.

**Table 3.** Participating entities and their endpoints.

Name	HTTP Port	Admin API Port	Webhook Port
Researcher	8030	8031	8032
IRB	8010	8011	8012
CDM Provider	8050	8051	8052
CDM Consumer	8060	8061	8062

ACA-py can also notify its controller when an event has occurred. It supports webhooks that allow immediately obtaining an update of what happened. Requests and responses between controllers configured through ACA-py are transmitted as HTTP requests, and webhook notifications are delivered as a result of processing. Webhook is an asynchronous HTTP callback on an event occurrence. It is a simple server-to-server communication for reporting a specific event occurred on a server. The server on which the event occurred will fire an HTTP POST request to another server on a URL that is provided by the receiving server.

In this paper, each of the cloud CDM subjects operates their own agents acting as a peer, and transactions between peers are maintained in a distributed ledger. Agent-to-agent communication is based on the DiDComm specification to support bilateral communication through a trusted channel.

#### 4.2. Experimental Result

The simulation environment setup starts with the registration of the entity researcher named Alice on each IRB. To establish the connection between IRB and Alice, IRB advertises an invitation data, Alice delivers the invitation message to IRB, and IRB responds to the accept message associated with the invitation. For peer-to-peer communication, Aries Interop Profile (AIP) uses 20. AIP is used to establish a connection between agents, exchange identity certificates, and perform transmission data through command delivery. After the identity is verified, the user's CDM data credential is performed.

After processing the registration information, the IRB sends a unique connection invitation message to Alice, as represented in Figure 10. The connection request message is used to communicate the DID document of the invitee (Alice) to the inviter (IRB). The @type attribute is a required string value that denotes that the received message is a connection request. After receiving the connection request, IRB evaluates the provided DID and DID Doc according to the DID Method Spec.

```
{"@type": "did:sov:BzCbsNYhMrjHiqZDTUASHg;spec/connections/1.0/invitation",
  "@id": "609089da-5224-4bf2-8093-e36c21221969",
  "serviceEndpoint": "http://192.168.65.3:8010",
  "label": "irb.agent",
  "recipientKeys": ["2mbtVTnEwDKZwW4AMqyFDnnTrNXXWgLFr2GszqNjISn"]
}
```

**Figure 10.** JSON format of IRB invitation attribute.

When IRB and researcher agents want to connect with each other, they establish a connection by DIComm, a series of messages that go back and forth to establish a connection and exchange information. In Figure 11, connection\_id is used to send a message between two agents.

```
{
  "invitation_mode": "once",
  "request_id": "60e9b302-5e73-4b26-b63b-2fa108a0617d",
  "rfc23_state": "request-sent",
  "created_at": "2021-07-30 06:21:48.258256Z",
  "updated_at": "2021-07-30 06:21:48.284719Z",
  "connection_id": "af76eeca-c877-45f1-bfea-531b98cc40cd",
  "their_role": "inviter",
  "my DID": "WceusqQ5hmeFkSGbizx8b8",
  "their_label": "irb.agent",
  "connection_protocol": "connections/1.0",
  "invitation_key": "2mbtVTnEwDKZwW4AMqyFDnnTrNXXWgLFr2GszqNjISn",
  "state": "request",
  "invitation_msg_id": "609089da-5224-4bf2-8093-e36c21221969",
  "routing_state": "none",
  "accept": "auto"
}
```

Figure 11. JSON format of the accepted message associated with the invitation.

In answer to the connect invitation, the IRB issues and offers a researcher a VC, represented in Figure 12 (segment of the issued credential), to be used to prove his/her identity when connecting to CDM provider. The VC is issued according to its schema definition in Figure 6. The credential is stored in the wallet of the researcher. The credential is generated based on IRB records including IRB number, name, affiliation, the status of GCP, etc. GCP stands for good clinical practice. This means that the clinical studies using CDM satisfy the clinical trial management criteria through the IRB.

```
{
  "referent": "7f0a809c-1e47-48c9-9921-6905db1152ff",
  "attrs": {
    "IRB_no": "2021-0008",
    "name": "Alice Smith",
    "timestamp": "1627626483",
    "role": "PI",
    "affiliation": "CNUH",
    "GCP": "1",
    "approved_date": "2021-03-28"
  },
  "schema_id": "BWvLd6qtFTQANsGbLDNSgQ:2:IRB schema:98.19.46",
  "cred_def_id": "BWvLd6qtFTQANsGbLDNSgQ:3:CL:42:irb.agent.IRB_schema",
  "rev_reg_id": null,
  "cred_rev_id": null
}
```

Figure 12. Researcher VC offered by the IRB upon registration.

Similarly, using CDM provider VC schema, the same setup is performed for the CDM provider. They aim to identify the CDM providers and the IRB issues and provide the CDM with a VC to allow the research to verify the CDM provider. Upon receiving the accessing CDM request, the CDM provider requires the researcher to present a valid verifiable credential (issued by the IRB), containing GCP in the allowed status of the credential in Figure 13. In the response to the CDM provider, the researcher presents a valid VC, with the allowed GCP granting the researcher permissions to access CDM data in that CDM provider. As shown in Figure 13, the result of the process is handled by the researcher. The

proof from the researcher is validated by CDM provider. Using AnnoCreds, the validation process is based on the GCP attribute in VC.

```

"affiliation": "XXXX",
"GCP": "1",
"approved_date": "2021-03-28"

# Request proof of IRB from alice researcher
IRB | Presentation: state = request-sent, pres_ex_id = b74f5cb7-05af-4d1a-b009-2d548d6ee6eb
IRB | Presentation: state = presentation-received, pres_ex_id = b74f5cb7-05af-4d1a-b009-2d548d6ee6eb

# Process the proof provided by X
# Check if proof is valid
IRB | Presentation: state = done, pres_ex_id = b74f5cb7-05af-4d1a-b009-2d548d6ee6eb
IRB | Proof = true

```

**Figure 13.** The result of the process the proof by using ZKP.

Figure 14 shows a proof, which is part of the credential issued by IRB, provided by the researcher to IRB and the CDM provider showing that the researcher is qualified. IRB verifies the qualification in the ZKP method based on the properties of the provided proof. Using the proof IRB, IRB give a permission of the CDM data request qualification when the attribute value of GCP is greater than 0.

```

{"indy": {"requested_predicates": {"0_age_GE_uuid": {"cred_id": "f39cfe6f-a35e-4e80-b0d6-aa281b00d90f"}, "requested_attributes": {"0_name_uuid": {"cred_id": "f39cfe6f-a35e-4e80-b0d6-aa281b00d90f", "revealed": true}, "0_date_uuid": {"cred_id": "f39cfe6f-a35e-4e80-b0d6-aa281b00d90f", "revealed": true}, "0_degree_uuid": {"cred_id": "f39cfe6f-a35e-4e80-b0d6-aa281b00d90f", "revealed": true}}, "self_attested_attributes": {}}}

```

**Figure 14.** The proof.

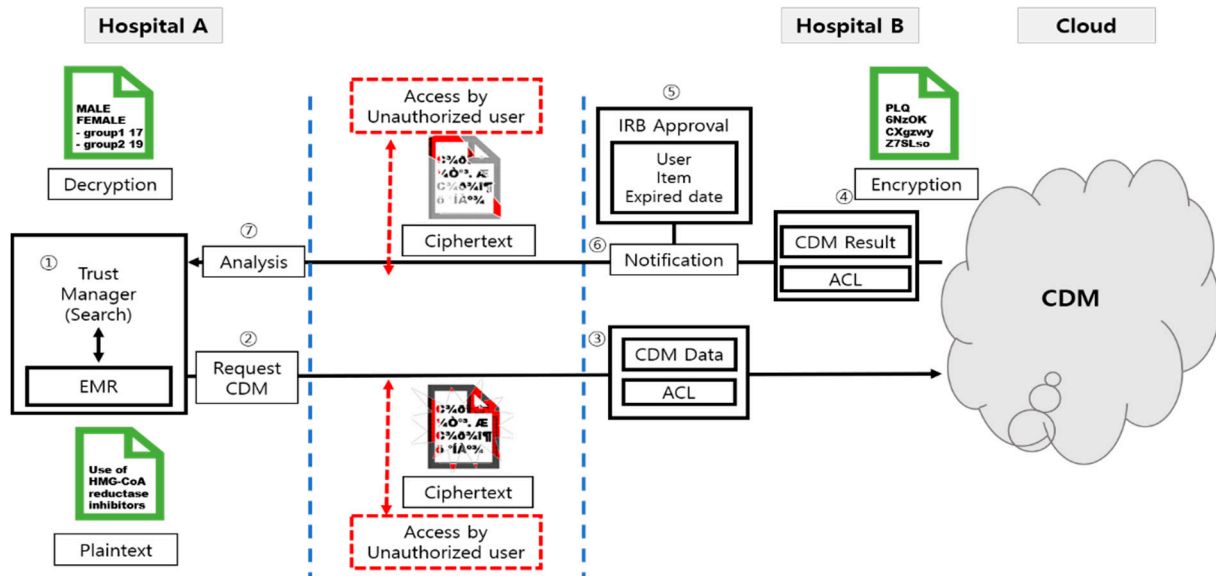
#### 4.3. Discussion

Privacy, security, and usability: the healthcare data are sensitive by nature, and they need a maximum of security against data breaches and privacy disclosure when exchanging the data, especially after enabling third parties' medical services to interact with the system. Medical data formats such as CDM for joint use have been developed for the participation of multiple hospitals and research institutions, and a stronger response method that is not vulnerable to security is needed. In order to further improve usability, in this paper, reliable cloud CDM research is conducted using DID based on blockchain.

In the construction, operation, and utilization of CDM, it is generally used only in the computer network within individual hospitals so that it is maintained at the same security level as general medical information. However, the problem of information leakage may occur due to insufficient systems or regulations to take responsibility for information security and prepare countermeasures in multi-institutional combined research. In addition, although CDM is mainly built on a cloud-based basis, security for conversion and conversion and de-identification of personal information in the hospital information system cannot be performed by building a clear solution or system. Instead, CDM is verified by the business procedure to confirm or pledge not to leak personal information by the programmer and system manager who performs the conversion and has a very weak structure. Therefore, clinical information in hospitals usually has to go through the consent of the patient who is the data subject and approval by the IRB. In addition, there is a restriction that researchers must use medical data only inside the hospital.

Figure 15 shows the flow of access control in cloud CDM. When a manager with authority sends a plaintext inquiry to the CDM (①②), the access control list and CDM data are transmitted to the cloud CDM (③). The cloud CDM performs the following detailed steps

and then sends the encrypted request result and ACL. The user, data approval range, period of use, etc., are subject to IRB review, and if approved (④⑤⑥), the user finally performs the analysis with the CDM result value (⑦). During a series of processes, data are encrypted, and unauthorized users' access is blocked so that the contents cannot be checked.



**Figure 15.** Flow of access control in cloud CDM. ① Search in trust manager ② Request data from the hospital where the data are available ③ Request for CDM data, attachment of access control list ④ Request result, ACL ⑤ IRB approval (user, data approval range, period of use) ⑥ Approval notice ⑦ User analysis.

## 5. Conclusions

Some businesses, including those that analyze CDM in public health research, which deals with sensitive information, may require a certain level of privacy and security. CDM for data sharing and utilization of medical institutions requires access to various patient medical information. It is used for disease research and customized medical care. Intrinsicly the CDM data are highly sensitive, and they need maximum security against data breaches and privacy disclosure when exchanging data. The cloud CDM provides interoperability for the participation of multiple hospitals and serves as an information-based study for customized and user-centered healthcare. However, reliable management of safe and transparent medical information of personal information is required.

The cloud CDM proposed applies DID and blockchain technology for secure access control that occurs when a researcher accesses it. The proposed service model is used to provide the credential of the researcher in the process of creating and accessing the CDM data of the designed secure cloud CDM. It does not consider the interaction with the existing system for establishing the initial trustiness of entities participating in the cloud CDM and suggests showing that the DID is used as a method for identification.

The prototype is an extension of the delivery of encrypted CDM using DID and describes the identification by limiting the use case of the CDM data of the researcher registered in the cloud CDM. This proposed method aspires to provide a unified and efficient data access control policy management framework. The designed model was verified by applying the ophthalmic CDM data of domestic hospitals. It provides strong security and ensures both the integrity and the availability of CDM data.

**Author Contributions:** Conceptualization, Y.B.P. and Y.K.; methodology, Y.B.P.; software, Y.K.; validation, Y.K., Y.B.P. and J.C.; formal analysis, Y.K.; investigation, J.C.; resources, Y.B.P.; data curation, Y.B.P.; writing—original draft preparation, Y.K.; writing—review and editing, J.C.; visualization, Y.K.; supervision, J.C.; project administration, J.C.; funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Korea Environmental Industry & Technology Institute (KEITI), grant number RE202101551 and The APC was funded by Ministry of Environment (ME).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** This work was supported Korea Environmental Industry & Technology Institute (KEITI) grant funded by the Korea government (Ministry of Environment). Project No. RE202101551, the development of IoT-based technology for collecting and managing big data on environmental hazards and health effects.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Shivade, C.; Raghavan, P.; Fosler-Lussier, E.; Embi, P.J.; Elhadad, N.; Johnson, S.B.; Lai, A.M. A review of approaches to identifying patient phenotype cohorts using electronic health records. *J. Am. Med. Inform. Assoc.* **2014**, *21*, 221–230. [[CrossRef](#)] [[PubMed](#)]
2. Ferreira, J.C.; Ferreira da Silva, C.; Martins, J.P. Roaming service for electric vehicle charging using blockchain-based digital identity. *Energies* **2021**, *14*, 1686. [[CrossRef](#)]
3. Liu, B.; Yuan, X.-T.; Yu, Y.; Liu, Q.; Metaxas, D. Decentralized Robust Subspace Clustering. *Proc. AAAI Conf. Artif. Intell.* **2016**, *30*, 3539–3545. Available online: <https://ojs.aaai.org/index.php/AAAI/article/view/10473> (accessed on 1 July 2021).
4. Xia, S.; Zheng, S.; Wang, G.; Gao, X.; Wang, B. Granular ball sampling for noisy label classification or imbalanced classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**. [[CrossRef](#)]
5. You, S.C.; Lee, S.; Cho, S.Y.; Park, H.; Jung, S.; Cho, J.; Yoon, D.; Park, R.W. Conversion of National Health Insurance Service-National Sample Cohort (NHIS-NSC) database into observational medical outcomes partnership-common data model (OMOP-CDM). *Stud. Health Technol. Inf.* **2017**, *245*, 467–470.
6. Chadwick, D.W. Federated identity management. Foundations of security analysis and design v. *Lect. Notes Comput. Sci.* **2009**, *5705*, 96–120.
7. Jayaraman, I.; Mohammed, M. Secure Privacy Conserving Provable Data Possession (SPC-PDP) framework. *Inf. Syst. E-Bus. Manag.* **2019**, *1*–27. [[CrossRef](#)]
8. Xiong, L.; Li, F.G.; Zeng, S.K.; Peng, T.; Liu, Z.C. A Blockchain-based privacy-awareness authentication scheme with efficient revocation for multi-server architectures. *IEEE Access* **2019**, *7*, 125840–125853. [[CrossRef](#)]
9. Cho, J.H.; Kang, Y.; Park, Y.B. Secure delivery scheme of common data model for decentralized cloud platforms. *Appl. Sci.* **2020**, *10*, 7134. [[CrossRef](#)]
10. Pănescu, A.T.; Manta, V. Smart contracts for research data rights management over the ethereum blockchain network. *Sci. Technol. Libr.* **2018**, *37*, 235–245. [[CrossRef](#)]
11. Androulaki, E.; Barger, A.; Bortnikov, V.; Cachin, C.; Christidis, K.; de Caro, A.; Enyeart, D.; Ferris, C.; Laventman, G.; Manevich, Y.; et al. Hyperledger fabric: A distributed operating system for permissioned blockchains. In *Proceedings of the Thirteenth EuroSys Conference, Porto, Portugal, 23–26 April 2018*; ACM: New York, NY, USA, 2018; p. 30.
12. Dagher, G.G.; Mohler, J.; Milojkovic, M.; Marella, P.B. Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. *Sustain. Cities Soc.* **2018**, *39*, 283–297. [[CrossRef](#)]
13. Silberschatz, A.; Korth, H.F.; Sudarshan, S. *Database System Concepts*; McGraw-Hill: New York, NY, USA, 1997.
14. Hyperledger/Aries-Cloudagent-Python. Available online: <https://github.com/hyperledger/aries-cloudagent-python> (accessed on 1 April 2021).
15. Reed, D.; Sporny, M.; Longley, D.; Allen, C.; Grant, R.; Sabadell, M. Decentralized Identifiers (DIDs) v1.0—Core Architecture, Data Model, and Representations. IT Security and Privacy—A Framework for Identity Management (ISO/IEC 24760-1). Available online: <https://www.w3.org/TR/did-core/> (accessed on 1 March 2021).
16. Blumenthal, D.; Tavenner, M. The “meaningful use” regulation for electronic health records. *N. Engl. J. Med.* **2010**, *363*, 501–504. [[CrossRef](#)]
17. Jensen, P.B.; Jensen, L.J.; Brunak, S. Mining electronic health records: Towards better research applications and clinical care. *Nat. Rev. Genet.* **2012**, *13*, 395–405. [[CrossRef](#)] [[PubMed](#)]
18. Glicksberg, B.S.; Oskotsky, B.; Giangreco, N.; Thangaraj, P.M.; Rudrapatna, V.; Datta, D.; Butte, A.J. ROMOP: A light-weight R package for interfacing with OMOP-formatted electronic health record data. *JAMIA Open* **2019**, *2*, 10–14. [[CrossRef](#)] [[PubMed](#)]
19. Reps, J.M.; Schuemie, M.J.; Suchard, M.A.; Ryan, P.B.; Rijnbeek, P.R. Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. *J. Am. Med. Inf. Assoc.* **2018**, *25*, 969–975. [[CrossRef](#)] [[PubMed](#)]
20. Voss, E.A.; Makadia, R.; Matcho, A.; Ma, Q.; Knoll, C.; Schuemie, M.; Ryan, P.B. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J. Am. Med. Inf. Assoc.* **2015**, *22*, 553–564. [[CrossRef](#)]
21. Garza, M.; Del Fiore, G.; Tenenbaum, J.; Walden, A.; Zozus, M.N. Evaluating common data models for use with a longitudinal community registry. *J. Biomed. Inform.* **2016**, *64*, 333–341. [[CrossRef](#)]



22. Hripcsak, G.; Duke, J.D.; Shah, N.H.; Reich, C.G.; Huser, V.; Schuemie, M.J.; Ryan, P.B. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for observational researchers. *Stud. Health Technol. Inform.* **2015**, *216*, 574.
23. Yoon, D.; Ahn, E.K.; Park, M.Y.; Cho, S.Y.; Ryan, P.; Schuemie, M.J.; Park, R.W. Conversion and data quality assessment of electronic health record data at a Korean tertiary teaching hospital to a common data model for distributed network research. *Healthc. Inform. Res.* **2016**, *22*, 54–58. [[CrossRef](#)] [[PubMed](#)]
24. Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system. *Decent. Bus. Rev.* **2008**, 21260–21268.
25. Alamri, B.; Javed, I.T.; Margaria, T. A GDPR-compliant framework for IoT-based personal health records using blockchain. In Proceedings of the 2021 11th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Paris, France, 19–21 April 2021; pp. 1–5.
26. Simply Vital Health. Available online: <https://www.simplyvitalhealth.com/> (accessed on 29 December 2018).
27. Roehrs, A.; da Costa, C.A.; da Rosa Righi, R. OmniPHR: A distributed architecture model to integrate personal health records. *J. Biomed. Inform.* **2017**, *71*, 70–81. [[CrossRef](#)]
28. Landau, S.; Le van Gong, H.; Wilton, R. Achieving privacy in a federated identity management system. In *Financial Cryptography and Data; Dingledine, R., Golle, P., Eds.; Security 2009. Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2009; Volume 5628. [[CrossRef](#)]*
29. Allen, C. The Path to Self-Sovereign Identity. Life with Alacrity. Available online: <http://www.lifewithalacrity.com/2016/04/the-path-to-self-sovereign-identity.html> (accessed on 1 July 2021).
30. Hardjono, T.; Pentland, A. Verifiable anonymous identities and access control in permissioned blockchains. *arXiv* **2019**, arXiv:1903.04584.
31. Shrestha, A.K.; Vassileva, J. Blockchain-based research data sharing framework for incentivizing the data owners. In Proceedings of the International Conference on Blockchain, Seattle, WA, USA, 25–30 June 2018; Lecture Notes in Computer Science. Springer: Berlin/Heidelberg, Germany, 2018; Volume 10974, pp. 259–266.
32. Augot, D.; Chabanne, H.; Chenevier, T.; George, W.; Lambert, L.; Augot, D.; Chabanne, H.; Chenevier, T.; George, W.; Lambert, L. A user-centric system for verified identities on the Bitcoin blockchain. In *Data Privacy Management, Cryptocurrencies and Blockchain Technology; Springer: Oslo, Norway, 2017; Volume 10436, pp. 390–407.*
33. Halpin, H. NEXLEAP: Decentralizing identity with privacy for secure messaging. In Proceedings of the 12th International Conference on Availability, Reliability and Security, Reggio Calabria, Italy, 29 August–1 September 2017; pp. 1–10.
34. Babkin, S.; Epishkina, A. Authentication protocols based on one-time passwords. In Proceedings of the 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), Saint Petersburg, Russia, 28–31 January 2019; pp. 1794–1798.
35. Zhang, R.; Xue, R.; Liu, L. Security and privacy on blockchain. *ACM Comput. Surv.* **2019**, *52*, 1–34. [[CrossRef](#)]
36. Taking the Sovrin Foundation to a Higher Level: Introducing SSI as a Universal Service. Available online: <https://sovrin.org/taking-the-sovrin-foundation-to-a-higher-level-introducing-ssi-as-a-universal-service/> (accessed on 10 August 2020).
37. Meralli, S. Privacy-preserving analytics for the securitization market: A zero-knowledge distributed ledger technology application. *Financ. Innov.* **2020**, *6*, 1–20. [[CrossRef](#)]